Class Notes for CS5229 Lecture 3

Semester 1, 2009/10

1 Introduction

The major differences in the derivation between this class and Padhye's [Padhye et al. 1998] are:

- I am using average values in the derivation. This is not rigorous, but we are building an approximation model anyway, and the approximated value will be accurate enough if the variance of these random variables are small. This is a common technique used in performance modeling.

- I assume that p is small and therefore, in several places, I am ignoring the terms p^k when $k > 1$. Furthermore, the term p^k for $k < 0$ is large and dominates other constant terms. We will ignore the constant terms in this case.

- I assume that TCP increases its window for every segment acknowledged instead of for every acknowledgment segments received (i.e., the parameter b in the paper is going to be 1 for this lecture).

2 Table of Notations

B	throughput of the connection
Y	number of packets sent in a triple-duplicate period
A	duration of a triple duplicate period
p	probability that a packet is lost given that it is the first packet in a round or the preceding packet in its round is not lost
α	the first packet lost in a triple-duplicate period (we order packets sent in a triple-duplicate period as 1, 2, 3, ...)
β	number of packets sent in the last round
W	the window size at the end of a triple-duplicate period (i.e., the window size when packet α is sent).
X	the round where the packet is lost
Q	probability that a loss indication ending a triple duplicate period is a timeout
R	number of packets sent during the timeout period
Z^{TO}	duration of a timeout period
T_O	value of the timeout timer

3 Triple-Duplicate ACKs Only

We begin by considering only packet losses signaled by a triple duplicate ACKs. We call the period between two consecutive triple duplicate events as a *triple-duplicate period*. Each period consists of a series of *round*. A round starts with the back-to-back transmission of one congestion-window worth of packets, and ends with the receipt of the first ACK for these packets.

We assume that if a packet is lost, then the following packets in the same round are lost as well. Assuming that such loss events are evenly distributed, then packet lost occurs on average every $1/p$ packet. In each triple-duplicate period, the α-th packet is lost, where

$$\alpha = \frac{1}{p}$$

If W is the congestion window at the time α is sent, we can send $W-1$ more packets after packet α, before triple duplicates are received. Therefore, the total number of packets sent during a triple-duplicate period, Y, is

$$\begin{aligned} Y &= W - 1 + \alpha \\ &= W + \frac{1-p}{p} \end{aligned} \tag{1}$$

Let X be the round number for packet α. The window size at the beginning of a triple-duplicate round is $W/2$, and at round X is W. Each round increases the window size by 1. Therefore,

$$\begin{aligned} W &= \frac{W}{2} + X \\ X &= \frac{W}{2} \end{aligned} \tag{2}$$

Chen et. al. [Chen et al. 2006] pointed out that W should be $W/2 + X - 1$. But we stick to Padhye's version here.

Furthermore, let β be the number of packets sent in the last round. Assuming that the number of packets sent in the last round is uniformly distributed between 1 and W, then

$$\beta = \frac{W}{2}$$

If we sum up the number of packets sent in each round, we get Y.

$$\begin{aligned} Y &= \beta + \sum_{k=0}^{k=X-1} \left(\frac{W}{2} + k\right) \\ &= \beta + \frac{X}{2}\left(\frac{W}{2} + W - 1\right) \\ &= \frac{W}{2} + \frac{W}{4}\left(\frac{W}{2} + W - 1\right) \\ &= \frac{3W^2}{8} + \frac{W}{4} \end{aligned} \tag{3}$$

From Equations (1) and (3), we get

$$
\begin{aligned}
W + \frac{1-p}{p} &= \frac{3W^2}{8} + \frac{W}{4} \\
\frac{3}{8}W^2 - \frac{3}{4}W - \frac{1-p}{p} &= 0 \\
3pW^2 - 6pW - 8(1-p) &= 0 \\
W &= \frac{6p + \sqrt{36p^2 - 96p(1-p)}}{6p} \\
W &= 1 + \sqrt{1 + \frac{8}{3}\left(\frac{1}{p} - 1\right)} \\
W &\approx \sqrt{\frac{8}{3p}} \qquad (4)
\end{aligned}
$$

Packet α is lost in the penultimate (second before last) round, therefore there are a total of $X+1$ rounds in a triple-duplicate period, and

$$
\begin{aligned}
A &= (X+1)RTT \\
&= \left(\frac{W}{2} + 1\right) RTT \\
&= \left(\frac{3}{2} + \sqrt{\frac{2}{3p} - \frac{5}{12}}\right) RTT \\
&\approx \sqrt{\frac{2}{3p}} RTT \qquad (5)
\end{aligned}
$$

The throughput can be computed as

$$
\begin{aligned}
B &= \frac{Y}{A} \\
&= \frac{\frac{1-p}{p} + W}{A} \\
&\approx \frac{\frac{1}{p} + \sqrt{\frac{8}{3p}}}{\sqrt{\frac{2}{3p}} RTT} \\
&\approx \frac{1}{RTT}\sqrt{\frac{3}{2p}} \qquad (6)
\end{aligned}
$$

4 Triple Duplicate ACKs and Timeouts

Consider the penultimate round in a triple duplicate period. W packets are sent. If $W > k$ and only k ($k < 3$) packets are ACKed, then we have a timeout. Let $A(W,k)$ be the probability that the first k packets are ACKed given that there is a sequence of one or more losses, then,

$$
\begin{aligned}
A(W,k) &= \frac{(1-p)^k p}{1-(1-p)^W} \\
&\approx \frac{(1-kp)p}{Wp} \\
&\approx \frac{1}{W}
\end{aligned}
$$

Other possibilities that lead to timeout exist but we assume that they are rare and ignore them in this approximation.

We can now estimate the probability of timeout as follows.

$$
\begin{aligned}
Q &= \begin{cases} 1 & \text{if } W \le 3 \\ \sum_{k=0}^{2} A(W,k) & \text{otherwise} \end{cases} \\
&= min\left(1, \frac{3}{W}\right) \qquad (7)
\end{aligned}
$$

Assume that one packet is sent during one timeout sequence. The number of packets sent during the timeout period would be the same as the number of timeouts. If the packet retransmitted after a timeout period is lost (with probability p), we get another timeout. The expected number of timeout R, is

$$
\begin{aligned}
R &= 1 + p + p^2 + ... \\
&\approx 1 + p
\end{aligned}
$$

The probability of more than two consecutive timeouts is too small to be considered here. Since TCP exponentially backs off the retransmission timer (i.e., doubles its value) for every consecutive timeout, the expected duration is approximately

$$
Z^{TO} \approx T_O(1+2p)
$$

We can now approximate the throughput considering both timeout and triple duplicates.

$$
\begin{aligned}
B &= \frac{Y + QR}{A + QZ^{TO}} \\
&= \frac{Y + min(1, \frac{3}{W})(1+p)}{A + min(1, \frac{3}{W})T_0(1+2p)} \\
&= \frac{\frac{1}{p} + \sqrt{\frac{8}{3p}} + min(1, 3\sqrt{\frac{3p}{8}})(1+p)}{RTT\sqrt{\frac{2}{3p}} + min(1, 3\sqrt{\frac{3p}{8}})T_O(1+2p)} \\
&= \frac{1 + \sqrt{\frac{8p}{3}} + min(1, 3\sqrt{\frac{3p}{8}})p(1+p)}{RTT\sqrt{\frac{2p}{3}} + min(1, 3\sqrt{\frac{3p}{8}})T_O p(1+2p)} \\
&\approx \frac{1}{RTT\sqrt{\frac{2p}{3}} + min(1, 3\sqrt{\frac{3p}{8}})T_O p(1+2p)}
\end{aligned}
$$

5 Impact of Window Limitation

Suppose the receiver advertises a maximum buffer size W_{max}, then the sender's congestion window cannot grow beyond W_{max}. Suppose the window can grow to W_{max} without causing any packet loss, then the throughput will simply be W_{max}/RTT.

Putting the equations together, we have a complete model:

$$
B \approx min\left(\frac{W_{max}}{RTT}, \frac{1}{RTT\sqrt{\frac{2p}{3}} + min(1, 3\sqrt{\frac{3p}{8}})T_O p(1+2p)}\right) \qquad (8)
$$

The original Padhye's equation is

$$
B \approx min\left(\frac{W_{max}}{RTT}, \frac{1}{RTT\sqrt{\frac{2p}{3}} + min(1, 3\sqrt{\frac{3p}{8}})T_O p(1+32p^2)}\right)
$$

Note the term $(1+32p^2)$ from Padhye's original equation is different from mine.

References

CHEN, Z., BU, T., AMMAR, M., AND TOWSLEY, D. 2006. Comments on “Modeling TCP Reno performance: a simple model and its empirical validation”. *IEEE/ACM Trans. on Networking 14*, 2, 451–453.

PADHYE, J., FIROIU, V., TOWSLEY, D., AND KUROSE, J. 1998. Modeling TCP throughput: a simple model and its empirical validation. *SIGCOMM Comput. Commun. Rev. 28*, 4, 303–314.